Krebs, W.K., Xing, J., Ahumada, A.J. Jr. (2002). A simple tool for predicting the readability of a monitor. Proceedings of the 46th Annual Meeting Human Factors and Ergonomics Society, 46.

A SIMPLE TOOL FOR PREDICTING READABILITY ON A MONITOR

[#&]William K. Krebs, [*]Jing Xing, and [*]Albert J. Ahumada, Jr.
[#]Federal Aviation Administration, Washington, D.C.
[&]Naval Research Laboratory, NR Unit 206, Washington, D.C. and
[*]NASA Ames Research Center, Moffett Field, CA

Background: Human factor practitioners are sometimes required to provide a rapid answer to the acquisition question, "How readable is text on this monitor?" Unfortunately, text readability is not listed on the manufacturer's brochure. This study's results were used to construct a simple tool to quickly assess text readability on a monitor based only on the luminance of the text and background colors. **Methods:** The text readability of three observers was measured for four colors (red, green, yellow, white), three intensity levels (.20, .25, .45), at three distances (1.62, 2.38, 3.16 meters) on a 20" color monitor. **Results:** The minimum error-free readable font size could be determined from the text/background luminance contrast alone. Thus luminance, not color, determined readability. From these results, a MATLAB program was developed that prompts for background and text RGB values and returns the minimal error-free readable font size. **Conclusions:** The tool fairly robustly and quickly assesses text readability on a monitor.

INTRODUCTION

Quite often, an acquisition program manager will request the assistance of a human factor professional to answer a specific question in a short period of time. The acquisition program manager's job performance is based on the ability to deliver the product on schedule. If the product is delayed, this will negatively impact the entire system as well as reflect poorly on the manager's ability to manage a program. Therefore, any delays are not tolerated. When human factors problems arise, the acquisition office wants a quick solution usually at that instant or within a day or two. Unfortunately, the human factors professional does not have the luxury to conduct a study to answer a specific problem that cannot always be answered by referring to a human factors reference. Therefore, to assist human factors professionals, simple and robust analytical tools should be developed to aid specific human factors problem-solving situations.

Today's cathode ray tube and liquid crystal display monitors allow an observer to choose between a wide range of functions and capabilities. There are numerous standards and recommended procedures to characterize the color and luminance of a display (Society of Automotive Engineers, 1989; Video Electronics Standards Association, 2001), color field uniformity (Electronic Industry Association, 1995), contrast (Electronic Industry Association, 1987a), specular gloss (Electronic Industry Association, 1987b), raster response (Electronic Industry Association, 1987c), resolution (Video Electronics Standards Association, 2001), and to measure the monitor's mechanical and physical characteristics (Video Electronics Standards Association, 2001). In addition, a monitor manufacture typically provides qualitative performance metrics about each monitor, i.e., the monitor's weight, intensity, frame rate, screen size, resolution, and addressability. However, it does not quantify text readability. Users often want to know what is the readability of the monitor for a given font size, text color, and background color. So far there is no simple method to determine a monitor's readability other than conducting time consuming and expensive human performance readability tests for specific types of text.

A common technique used by human factor professionals to assess monitor's readability is to measure observers' reading performance for displayed text. Text can be depicted by luminance contrast or color contrast. A number of studies have demonstrated that readability can be predicted from text contrast

(Ahumada, 1996, Scharff, Hill, and Ahumada, 2000). Legge and Rubin (1986) also showed that luminance contrast determined participants reading rate regardless of text color. Although it has been reported that either color or luminance contrast can support reading (Legge, Parish, Luebker, and Wurm, 1990), Knoblauch, Arditi, and Szlyk (1991) found that, although colors affected reading performance at threshold luminance contrast and at very small text size, the performance was unaffected by chromatic contrast in the presence of suprathreshold luminance contrast (0.12) over a large range of text font sizes.

Readability experiments consume resources and time. It would be nice to have a simple tool that would allow human factors professionals to predict text readability on a monitor by entering a small number of recorded photometric values from the monitor. The objective of this paper is to demonstrate that a simple software tool can be used to predict the readability of a monitor. We first performed human readability tests, varying factors that might affect text readability, such as luminance contrast, color, and viewing distance. We found the experimental results could be fit by an analytic equation that predicted the readability from the text size in visual angle and luminance contrast. Finally, we wrapped the equation in a software tool to assess text readability for color monitors. The simple tool, with a small number of photometric measurements as input to estimate a luminance response function, allows user to input any given background and text RGB values to determine minimal error-free readable font size. We used this tool to predict the minimal error-free readable font size for a proposed Federal Aviation Administration color replacement tower display monitor. The simple tool was also validated for another monitor to determine how well it could predict readability performance.

METHODS

Observers: Three observers (ages 34, 40, and 40) had normal or corrected-to-normal visual acuity and normal color vision as tested with the Farnsworth Dichotomous Test for Color Blindness and Dvorine Color Plate test. Observers had 14 years (σ = 3) of air traffic control experience. Informed consent was obtained from all observers. All observers were naïve to the experimental hypothesis.

Apparatus: Stimuli were displayed on a General Digital 20" AMLCD color monitor. Observers viewed the screen from three different distances of approximately 1.62, 2.38, and 3.16 meters for distances 1 through 3, respectively. The screen resolution was 1024 by 768 pixels with 3.2 pixels/mm in the vertical and horizontal direction.

Stimuli: Figure 1 illustrates the eye chart used in the experiment. The eye chart consisted of eight rows of letters (Lucida Console font) with each row containing nine unique letters that are commonly used in the Snellen eye chart. Physical x and y pixel dimensions of each character box from lowest to highest rows were 8 by 10, 9 by 11, 10 by 12, 11 by 15, 12 by 16, 14 by 18, 15 by 19, and 16 by 22, respectively. Text color was red, green, yellow, white or black.

Figure 1. Eye chart stimulus

RGB input	Y	x	y
red	110	0.621	0.344
green	259	0.323	0.57
blue	51	0.152	0.137

Table 1 shows the CIE luminance Y and the color chromaticity coordinates (x, y) for each RGB channel alone at full input.

The screen background was set to black (0, 0, 0) with a mean luminance 2.62 cd/m^2. Red, green, yellow, or white text were displayed at 20%, 25% and 45% intensity levels, defined as multipliers of the input digital values (e.g., a digital RGB input value of 255, 0, 0 with a 25% intensity level would have a digital RGB input of 64, 0, 0).

Procedure: The observers' task was to start at the top of the screen and read each row of letters. No feedback was provided but observers were encouraged to guess. The experimenter located outside the observer's field-of-view recorded the vocal responses. Observers were allowed periodic rest throughout the experimental session.

Thirty-six trials (3 intensity levels, 3 distances, and 4 colors) were presented to each subject. Distance and color were randomly assigned within each block of intensity trials.

RESULTS

RGB-Luminance Computation

We made a series of RGB-Luminance measurements in order to derive the RGB-luminance relationship for a given monitor. The data showed that no single gamma value could fit the whole luminance range. Therefore, a piecewise linear interpolation was used to compute luminance for any given RGB values.

Lr+Lmin is the luminance for red only (r,0,0);
Lg+Lmin is the luminance for green only (0,g,0);
Lb+Lmin is the luminance for blue only (0,0,b);
Lrgb is the luminance for RGB = (r,g,b); where

$$Lrgb=Lr+Lg+Lb+Lmin;$$

Lr was measured for r = 64, 128, and 255. To obtain Lr for an arbitrary r, log(Lr) was computed by linear interpolation between the two adjacent measured log values. Lg and Lb were obtained similarly. The luminance for RGB = (r,g,b) was the sum Lr+Lg+Lb+Lmin.

Viewing Distance and Viewing Angle

The readability measurement was made at three viewing distances (1.62, 2.38, 3.16 meters). We converted the text font sizes from pixels into visual angle (arc min). When the percentage of correctly read letters vs. angular font size was plotted for each viewing distance, the three curves superimposed, indicating that the absolute font size and the absolute viewing distance were not important.

Minimal Font Size For Error-Free Reading

For each of the four colors and three intensities tested, the percentage of correctly read letters was plotted against the angular font size. The percentage correct increased with font size. The minimal font size for error-free reading was defined as the smallest font size for 100% correct responses. The minimal font sizes for the 4x3 color-intensity combinations were thus determined.

Text Contrast Versus Minimal Font Size

We defined text contrast as (Lt-Lb)/Lb, where Lt is the text luminance and Lb is the background luminance. The text contrasts for the four tested colors and three intensities were computed by converting RGB values into luminance as above. The results are plotted in Figure 2. The vertical axis is the minimal font size for error-free reading and the horizontal axis is the text contrast. The solid curve is a least squares fit of the data by the equation:

$$fontSize = A+B \exp(-|contrast|/C),$$

where A=5.028, B=7.434, and C=0.6297.

The RMS error of fit for the above equation was 0.23 arc min. We also computed the RMS error of this equation for each text color separately. The errors were 0.26, 0.14, 0.22, and 0.14 arc min for white, yellow, red and green, respectively. Thus, the data obtained with different text colors were all well fit by the same equation, suggesting that the color was not a factor.

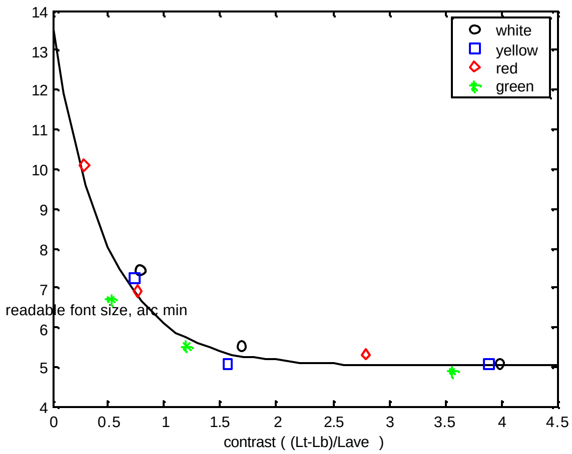

Figure 2. Readable size from luminance contrast function. The curve is a fitted equation. Different symbols represent the data for different text colors. The data for different text colors are fit well by the same function, showing that the color did not affect readability size.

MATLAB Tool to Predict Readability of a Monitor

The MATLAB program (available at http://www.hf.faa.gov/krebs/download.htm) prompts for background and text RGB values. It returns background and text luminance, text contrast and the minimal error-free readable font size (arc min). The screen data of a 20' CRT monitor were included in the RGB-Luminance tables within the program. Those RGB-luminance tables need to be revised to assess a new monitor.

MATLAB Tool Validation

Air traffic controllers' readability scores were obtained from another study (WJHTC Report, 2002) to determine how well the MATLAB tool could predict readability performance. Six air traffic controllers participated in a tower cab study to read data blocks from a flat panel color display during day (white background) and nighttime (black background) viewing conditions.

Table 2 shows the average percent of data blocks correctly read. Accuracy was higher with the black background. The distance by color by background interaction was significant. As distance from the screen increased, readability performance dropped significantly for all colors and backgrounds except red on black.

background	dist (m)	green	red	gray	yellow
black	1.21	94%	90%	91%	95%
black	2.13	88%	95%	83%	97%
black	3.05	67%	87%	81%	80%
white	1.21	94%	89%	78%	66%
white	2.13	82%	79%	83%	61%
white	3.05	58%	49%	52%	22%

Table 2. Average percentage of observer data blocks correctly read as a function of background, distance, and text color (data from WJHTC Report, 2002).

Table 3 shows predicted maximum error-free readability distances for a fixed text size (2.45 mm). The program was calibrated for the monitor used in the study. For example, 1.68 meters is predicted to be the greatest distance for accurately reading red text on a white background, because the smallest error free text was predicted to be 5.03 arc min in angular size. The predicted performance on the white background is similar to the behavioral scores, where yellow had the worse performance while white, green, and red were nearly equivalent. The program correctly predicts better performance on the black background, but predicts that the contrast is high enough that it does not limit reading acuity.

background	accur-acy	green	red	gray	yellow
black	100%	2.16	2.16	2.16	2.16
white	100%	1.80	1.68	1.91	1.40

Table 3. The MATLAB program prediction for the distance at which an observer can correctly read the text colors of Table 2 on the black and white backgrounds. Text size was assumed to be 2.45 mm.

CONCLUSIONS

Our readability experiments suggest that the minimal angular font size for error-free reading can be computed from the text luminance contrast, irrespective of text color and text resolution. In our experiments the crudest resolution was 8 by 10 pixels for a letter, so the predictions of our formula are only valid if the predicted size can be represented at this resolution or higher. A MATLAB program was developed to implement the formula from RGB values on a calibrated monitor. The program prompts for background and text RGB values and predicts the minimal error-free readable font size. The tool enables rapid prediction of text readability on a monitor.

REFERENCES

Ahumada, A.J. (1996). Simplified vision models for image quality assessment, SID International Digest of Technical Papers, 27, 397-400.

Electronic Industries Assocation (1987a). Contrast measurement of CRT's (TEPAC Publication No. TEP105-10). Arlington, VA: Global Engineering Documents.

Electronic Industries Assocation (1987b). Test method for specular gloss (TEPAC Publication No. TEP105-13). Arlington, VA: Global Engineering Documents.

Electronic Industries Assocation (1987c). Raster response measurement for monochrome cathode ray tubes (TEPAC Publication No. TEP105-08). Arlington, VA: Global Engineering Documents.

Electronic Industries Assocation (1995). Color field uniformity test procedures (TEPAC Publication No. TEP105-18). Arlington, VA: Global Engineering Documents.

Knoblauch K, Arditi A, Szlyk J. (1991) Effects of chromatic and luminance contrast on reading . Journal of the Optical Society of America, A, 8(2):428-39

Legge G.E. and Rubin, G.S. (1986). Psychophysics of reading. IV. Wavelength effects in normal and low vision. Journal of the Optical Society of America, A3, 40-51.

Legge G.E., Parish, D.H., Luebker, A., and Wurm, L.H. (1990). Psychophysics of reading. XI. Comparing luminance and color contrast. Journal of the Optical Society of America, A7, 2002-2010.

Scharff, L.V., Hill, A., Ahumada, A.J. (2000). Discriminability measures for predicting readability of text on textured backgrounds, Optics Express, 6(4), 81-90.

Society of Automotive Engineers (1989). Photometric and colorimetric measurement procedures for airborne direct view CRT displays (SAE Aerospace Recommended Practice ARP1782). Warrendale. PA.

Video Electronics Standards Association (2001). Flat panel display measurements standard: Version 2.0. Milpitas, CA: VESA.

William J. Hughes Technical Center (2002). Remote ARTS Color Display (R-ACD) human factors monitor readability assessment quick-look report. ACT-530 Unpublished Report: Atlantic City, New Jersey.

ACKNOWLEDGEMENTS

A special thanks to Kenneth Allendoerfer at the William J. Hughes Technical Center for his invaluable support in ensuring completion of this study. We also thank Dino Piccione for giving us the opportunity to participate in this study and his helpful critique of the paper. This work was partially supported by NASA Ames Research Center cooperative agreement NCC 2-1095 with the San Jose State University Foundation.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Federal Aviation Administration, United States Department of Transportation, nor the United States Government.